PCT/GB 2004 / 0 0 1 3 5 4

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

RECD 0 4 MAY 2004

WIPO PCT

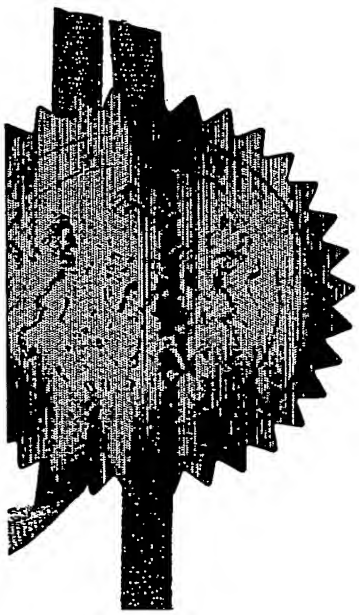
PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed M Cooke

Dated 27 April 2004

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

For Official use only

THE PATENT OFFICE
RN
- 7 APR 2003
RECEIVED BY FAX

07APR03 E798029-1 D10092
P01/7700 0.00-0307884.7

-7 APR 2003

Your reference **Patent K (UK)**

0307884.7

The **Patent Office**

Request for grant of a Patent

Form 1/77 Patents Act 1977

1 Title of invention

COMPUTER BASED SYSTEM FOR MANIPULATING DIGITAL MEDIA

2. Applicant's details

[X] First or only applicant

2a If applying as a corporate body: Corporate Name

Internet Pro Video Limited

Country

Great Britain

2b If applying as an individual or partnership
Surname

Forenames

2c Address

Mount Pleasant House
2 Mount Pleasant
Huntingdon Road
Cambridge

UK Postcode CB3 0RN

Country Great Britain

ADP Number 8604951001

☐ 2d	Second applicant (if any) Corporate Name Country
2e	Surname Forenames
2f	Address UK Postcode Country ADP Number
3	Address for service Agent's Name: Origin Limited Agent's Address: 52 Muswell Hill Road London Agent's postcode: N10 3JR Agent's ADP Number: C03274 -1270457002

4 Reference Number

Patent K (UK)

5 Claiming an earlier application date

An earlier filing date is claimed:

Yes ☐ No ☒

Number of earlier application or patent number

Filing date

15 (4) (Divisional)	8(3)	12(6)	37(4)
☐	☐	☐	☐

6 Declaration of priority

Country of filing	Priority Application Number	Filing Date

7 Inventorship

The applicant(s) are the sole inventors/joint inventors

Yes [ ] No [X]

8 Checklist

		Continuation sheets	
Claims	2	Description	6
Abstract	0	Drawings	0

Priority Documents	~~Yes~~/No
Translations of Priority Documents	~~Yes~~/No
Patents Form 7/77	~~Yes~~/No
Patents Form 9/77	Yes/~~No~~
Patents Form 10/77	~~Yes~~/No

9 Request

We request the grant of a patent on the basis of this application

Signed: Origin Limited. Date: 7 April 2003

(Origin Limited)

DUPLICATE

COMPUTER BASED SYSTEM FOR MANIPULATING DIGITAL MEDIA

Technical Field

This invention relates to a computer software system for manipulating digital media.

Background Art

Application software for editing digital video is an extremely sophisticated and powerful tool because it is primarily designed for, and sold to, the video professional. Such an individual requires access to many complex functions and is prepared to invest time and effort in learning to become skilled in their use. Historically, the terminology and conventions of Digital Editing have evolved from a traditional film editing environment where rushes are cut and spliced together to tell a story or follow a script. As digital mixer technology advanced new techniques were combined with these conventional methods to form the early pioneering software based digital editors.

To the video or film professional editing is second nature and the complexities of a time-based media go unnoticed since, having already grasped concepts and learned processes, they are able to concentrate on the nuances of different editing packages, of which there are many.

Conventionally these packages, through the use of a Graphical User Interface (GUI), attempt to provide an abstraction of the media in terms of many separate tracks of video and audio. These are represented on the output device in symbolic fashion and provision is made for interacting with these representations using an input device such as a mouse. Typically the purpose is to create a new piece of media as an output file, composed by assembling clips or segments of video and audio along a timeline that represents the temporal ordering of frames. Special effects such as wipes and fades can be incorporated, transparent overlays can be added, colour and contrast can be adjusted. The list of manipulations made possible by such tools is very long indeed. A typical system is described in, for example, Foreman; Kevin J., et. al, "Graphical user interface for a video editing system", U.S. Patent. 6,469,711.

It is possible, however, that an individual who is a consumer of media, rather than a producer, may need to perform a simple editing operation on a media file in order to accomplish their primary task; for example to give a multi-media presentation. In this case

such tools have their drawbacks. They may be too expensive to justify individually, or to have enough of in order to be available when or where needed. The limited amount of use and the small fraction of the capabilities used in such situations may make them uneconomic. The steep learning curve associated with such tools may mean that an inappropriate amount of effort is expended on something that is not the primary occupation or concern of the tool user. For occasional or infrequent use there will be reluctance on the part of any user repeatedly to switch environments or learn and relearn new tools to perform simple last minute tasks.

This situation parallels previous well-known situations where improvements in the availability, usability and price/performance ratio of consumer IT equipment, has caused a significant reappraisal of what is possible and a change in behaviour to exploit new possibilities. For example, the production of high-quality printed documents was once the province of highly skilled people using expensive and specialised equipment. Now anybody with a need to produce such a document, who has access to a computer and a word-processing program, can do so. A similar shift in paradigm is happening in Digital Video Editing, where there is a need for highly accessible and usable tools that focus on the needs of a new generation of user, and not necessarily try to recreate the feel of a traditional video editing environment.

Such tools must exhibit an intuitive, predictable and consistent behaviour as understood by a new generation of digital media professionals who may well be extremely familiar with the manipulation of documents of various kinds through a computer's GUI, but be completely unfamiliar with the characteristics of time-based media. The tools need not supercede long established and specialised tools used by trained professionals but, rather, provide a bridge in order that new users may be as comfortable working with time based media as they are working with documents.

Conventionally, video editors are structured as specialised 'monolithic' applications. Current software technology, however, is well capable of adding sophisticated editing functions to unrelated applications through the use of software 'plug-ins'. The Microsoft® DirectShow® Editing Services is an application programming interface (API) that is built on top of Microsoft® DirectShow® that allows video editing capabilities to be added to applications. In this example filters, implemented as Common Object Module components, are created and inter-connected to form filter graphs. As another example the QuickTime track based

architecture is the foundation of many modern day editors such as Adobe Premiere. It offers embedded API based access, resident below the application layer that provides for simple track manipulation. There is no reason why such plug-ins may not be deployed in applications that are not, primarily, designed for video editing.

Accordingly, these are the attributes of a tool that is more appropriate to the needs of such a consumer.

1. Simple and intuitive to use; in particular, little time and effort is required to learn enough to accomplish the task in hand.
2. Terminology and workflow consistent with a shift in convention towards action led digital media editing; e.g. 'Jog' to select as with modern VCR's and a simple crop ability to trim the running length of a piece of media.
3. Available whenever and wherever needed, even if the user did not foresee the need for such a tool until that need cropped up, i.e., the edit capability is provided as an intrinsic part of the environment by way of any player of the media.
4. Provides a consistent interface to the user irrespective of the type of 'container' application it is associated with. It would look exactly the same whether incorporated into an electronic text document, a spreadsheet, or a slide presentation.
5. Persistence of modifications; e.g. the user opens media object, or a document containing an embedded media object and expects any changes they make to persist between sessions.

Disclosure of Invention

The invention relates to a method for adding the capability for media-manipulation to software media players, such that this capability is intrinsic to the media player, and which comprises a set of tools that ensure that consistent behavioural, visual and functional aspects are maintained between media player applications.

Briefly, the invention works as follows.

According to one aspect of the invention a Graphical User Interface (GUI) for editing is implemented.

In the preferred embodiment of this aspect of the invention a plug-in module is loaded into the computer's memory to provide the specific functionality required. This software module has interfaces to a media delivery subsystem, such as the Microsoft® DirectShow® architecture for the Microsoft® Windows® platform that provides services for streaming, buffering, synchronisation, decoding and rendering of video and audio. Media is streamed into a local cache that provides for fine-grain scrubbing 'jog' and 'looping' of short sections around in and out points. A set of instructions is devised for each piece of media and its interaction with a timeline. Specific elements are constructed in memory to process these instructions and subsequently handle the media in a suitable form, as compatible with the media play architecture in operation. New and modified elements maybe constructed and reconstructed as required, each element may process but is not limited to a single set of instructions or piece of media.

The functionality provided by this software module consists of:-

- Graphics rendering to allow combination and/or overlay of graphical data for the GUI with pixels that are decoded from the video part of the media file and rendered into the video area on the screen.

- A cache for portions of the media file in the memory of the client machine.

- A state machine, whose transitions guide a user through a sequence of interactions with a graphical user interface (GUI).

- Graphics that implements visual feedback of the current state to the user.

- Graphics that implements a visual metaphor that provides the user with an intuitive understanding of the operation of the interface.

- An exporter for the persistence of the chosen manipulations, for example; to "Save" the processed media to memory creating a new media object or create a new set of instructions that describes the precise operation required to effect the manipulations for playback. Including but not limited to, such instruction as references to sections within a remote piece of media.

- Graphics that allows labels of various types to be added to significant parts of the media file in order to identify them as such.

- Graphics that allows the definition of actions to be taken when significant parts of the media file are encountered in normal playback.

In this embodiment of this aspect of the invention the GUI is provided by modules within the software framework that implements the media player, by the addition of visible user interface components (buttons, text boxes, etc.) either overlaid or actually burnt into the rendered video window (i.e. the pixels written to the framestore by the video renderer are overwritten). In the Windows Media® architecture, where software filter graph components are linked together to implement a media player, this functionality may be added into a video renderer filter or an overlay filter.

In another embodiment of this aspect of the invention the GUI is provided by software modules, other than those embedded within the media player framework, such as ActiveX controls.

According to another aspect of the invention elements are exchanged between instances of a media player.

In the preferred embodiment of this aspect of the invention the Windows® Media environment is employed such that one instance of the player may be used to manage the "master" timeline, while another allows clips to be trimmed to the desired length and then dragged and dropped into the "master" player instance. At this time the recipient instance may chose to combine the filter graph for the new piece of media with those already in existence, or it may chose to reconstruct a new filter graph based on the complexity and required interaction of the current timeline objects.

According to another aspect of the invention a process flow is provided that provides for untrained users to achieve their goal with minimum effort, and distraction from their primary task.

In the preferred embodiment of this aspect of the invention state machines help walk the users through operations to avoid mistakes and distil the complexity of editing into bounded

and easy to understand processes. Visual and tactile feedback will provide rapid confidence in the task and aid progress; e.g. to slim down a media object the user will select a "Start Here" in point and guided towards a "Stop Here" out point.

Effective confirmation methods are employed to inform and protect the actions of the user, visual metaphors will be provided from the embedded editor level to identify nodes of the current state machine. For example, the video window may show a filmstrip with the current frame highlighted subsequent frames normal and previous frames, i.e. those cropped, with a strike out marker.

According to another aspect of the invention meta-data in the media file is recognised by the system and used as a stream of control information that is used to assist editing operations.

In one embodiment of this aspect of the invention the meta-data may include but is not limited to:

- Timecode,
- Closed caption
- Edit points used during the creation of the media,
- Format-dependent properties such as GOP boundaries in MPEG,
- Data generated as a result of post-processing such as shot change information.

The control information identifies significant points in the media and triggers events that cause instructional or informative information to be displayed. For example dialogue boxes may pop up during playback with labels such as "Start Here" (IN) or "Stop Here" (OUT).

Industrial Applicability

As a simple example of the use of the invention consider this scenario. An education professional is preparing materials for a lecture they are about to give. A part of this is an electronic slide presentation with some of the slides containing embedded media objects. On running through the presentation they realise that it over-runs the time allowance if all the media clips are played in their entirety. Using the system described here the media can be quickly and efficiently trimmed to suit requirements, without the need for switching environments or applications.

CLAIMS

1. A computer based system for manipulating digital media, the system adding the capability for media-manipulation to software media players, such that this capability is intrinsic to the media player, and which comprises a set of tools that ensure that consistent behavioural, visual and functional aspects are maintained between media player based applications.

2. The system of claim 1 where media includes, but is not confined to, video and audio.

3. The system of claims 1 & 2 where manipulation includes, but is not confined to, the operations of one or more of:

   (a) Editing; trimming; annotating; effects; transitions; appearance; presentation;

4. The system of claims 1 - 3 comprising one or more of:

   (a) A software component that implements a cache for portions of the media file in the memory of the client machine.

   (b) A software component that implements a process equivalent to a state machine, whose transitions guide a user through a sequence of interactions with a graphical user interface (GUI).

   (c) A software graphics component of the GUI, that implements visual feedback to the user of the current state.

   (d) A software graphics component of the GUI that implements a visual metaphor that provides the user with an intuitive understanding of the operation of the interface.

   (e) A software graphics renderer component that allows combination and/or overlay of graphical data for the GUI with pixels that are decoded from the video part of the media file and rendered into the video area on the screen.

   (f) A software component that implements an export of the processed media to memory.

(g) A software component that implements the ability to read a description file(s) and construct playback in accordance with set instructions, or write such instructions from a current playback.

(h) A software component of the GUI that allows labels or triggers of various types to be added to significant parts of the media file in order to identify them as such.

5. The system of claims 1 - 4 where labels include, but are not confined to, edit in and out-points, shot boundaries, Group-Of-Picture (GOP) boundaries, closed-caption text and timecode.

6. The system of claims 1 - 5 where the triggers include, but are not confined to, initiate pop-ups, hold frames for a given duration, loop and messaging.

7. The system of claims 1 - 5 additionally comprising:

A software decoder component that maps meta-data contained in the media file to labels, where the meta-data includes, but is not confined to, shot boundaries, Group-Of-Picture (GOP) boundaries, closed-caption and timecode.

8. The system of claims 1 - 7 additionally comprising:

A software agent component that maps aspects of the interactive behaviour of the user into configuration information that may modify aspects of the behaviour of the GUI.

9. The system of claims 1 - 8 additionally comprising:

A media file that may optionally be selected and played by the user, which provides instruction in the use of the GUI.